International Journal of Applied Mathematical Sciences,
ISSN 0973-0176 Vol.2 No.1(2005), pp. 31-38

http://www.gbspublisher.com/ijams.htm

A Note on Bilevel Optimization Problems

Pu-yan Nie
Department of Mathematics,
Jinan University,
Guangzhou, 510632, China,
Email: tpynie@jnu.edu.cn

Abstract

Stackelberg games, with which it is very successful in such fields as economics, management, politics and behavioral sciences, can be modelled as a bilevel optimization problem. In a bilevel programming problem, there is a hierarchy of decision makers which is called as leader and followers, respectively. There exist extensive literatures about static bilevel optimization problem if the followers act independently. When the followers act dependently, namely, a follower's action depend not only on the decisions of the leader but also on the responses from other followers, the lower level problem is a multi-objective programming problem. In this paper, we consider the bilevel programming problems where lower level problem is a multi-objective programming problem. Weighting methods are employed to analyze constraint conditions, namely, multi-objective programming problems. Some properties about mathematical program with multi-objective constraints are obtained in this paper.

AMS Subject Classification (2000): 90C26, 90C30, 90C31.

Key Words: Bilevel programming problems, Stackelberg games, multi-objective programming, Pareto optimal solution.

1. Introduction

The most fundamental form of bilevel programming problems (BPP) issues from the Stackelberg games [8]. Bilevel optimization plays exceedingly important roles in optimization fields. We revisit BPP but give a new problem in this work. The following problem is considered:

$$\begin{array}{ll} \min_{x,y} & F(x,y) \\ \text{subject to} & G(x,y) \leq 0, \\ & y \in \arg\min\{f(x,y) : g(x,y) \leq 0\}, \end{array} \tag{1}$$

where $x \in \mathcal{X} \subseteq R^n, y \in \mathcal{Y} \subseteq R^m$ and $F : R^{n+m} \to R, f : R^{n+m} \to R^N, G : R^{n+m} \to R^{m_0}, g : R^{n+m} \to R^{m_1}$ are all continuously differentiable. We call x, y as the variables of leader and followers, respectively. In previous BPPs, the followers act independently. In the problem (1), the followers interact when they make decisions. Actually, these models are very popular in practice. When we model the practical problems, the lower level problem is a multi-objective problem. We call this problems mathematical programming with multi-objective constraints or MPMOC in brief.

When the followers act independently, there exist extensive researches [7, 9, 10] and the cited references about BPPs. Under certain conditions, a bilevel optimization problem can be reformulated as a mathematical program with equilibrium constraints (MPEC) problem [3], which has recently drawn much attention in the optimization community [4, 6].

We give some related notation to (1). The following problem to (1), which is a multi-objective programming problem in the problem (1), is called lower level problem

$$\begin{array}{ll} \min_y & f(x,y) \\ \text{subject to} & g(x,y) \leq 0. \end{array} \tag{2}$$

And $\mathcal{F} = \{(x,y)|G(x,y) \leq 0, g(x,y) \leq 0, \}$ is called relaxed feasible region.

MPMOC is very important in practice. In Stackelberg games, it is useful investigating MPMOC to explain and to predicate some social phenomena and individual or group behaviors. We illustrate MPMOC with a simple example as follows:

Example 1 (Agent Problem) Some agency relations are Stackelberg games. Assume that there are N agents and a corporation and the agents are competitive or non-competitive. Let the cost function of corporation be $B(x,y)$ and the constraint be $G(x,y) \leq 0$. The cost function of the ith agent is $b^i(x,y)$ and the ith follower subjects constraints $g_i(x,y) \leq 0$. Then, the model is

$$\begin{array}{ll} \min_{x,y} & B(x,y) \\ \text{subject to} & G(x,y) \leq 0, \\ & y \in \arg\min_x \{(b^1(x,y), b^2(x,y), \cdots, b^N(x,y)) : \\ & g^i(x,y) \leq 0 \quad i = 1, 2, \cdots, N\}. \end{array}$$

This paper is organized as follows: In the next section, some definitions are presented for MPMOC. Some properties about MPMOC are given in the section 3. Remarks and future researches are proposed in the final section.

2. Some Definitions to MPMOC

For lower level problem, it is a multi-objective programming problem. There are many kinds of solutions to multi-objective programming problems, for example, admissible solution set, dominant strategies equilibrium and so on. Among them, Pareto optimal solutions are very important, which is listed as follows:

Definition 0.1 *For any x, the Pareto solution to x is*

$$P(x) = \{y | g(x, y) \le 0 \text{ and there exist no } \bar{y} \text{ such that } g(x, \bar{y}) \le 0, \quad f(x, \bar{y}) \preceq f(x, y)\}, \tag{3}$$

where $f(x, \bar{y}) \preceq f(x, y)$, which is named domination, means $f_i(x, \bar{y}) \le f_i(x, y)$ for all $i = 1, 2, \cdots, N$ and there exists at least a strict inequality.

We now aim to define the solutions to MPMOC and the solutions are determined by the idea of the decisions. Pareto solutions reflect the decisions of the followers. For leader, there exist all kinds of solutions. Some definitions related to (1) are presented as follows

Definition 0.2 *We call $x^\star$ a risk optimal decision if and only if*

$$x^\star \in \arg\min\{\min_{y \in P(x)} F(x, y)\}. \tag{4}$$

Definition 0.3 *We call $x^\star$ a conservative optimal decision if and only if*

$$x^\star \in \arg\min\{\max_{y \in P(x)} F(x, y)\}. \tag{5}$$

Definition 0.4 *We call $x^\star$ a mean optimal decision if and only if*

$$x^\star \in \arg\min\{\frac{\int_{y \in P(x)} F(x, y) dy}{\int_{y \in P(x)} dy}\}. \tag{6}$$

When there exists exact a solution to the lower level problem for any x, there are no differences about all kind of solution. When there are multiple solution to the lower level problem, all kinds of solutions are different. We point out that above three definitions relate close to the styles of the decisions of the leaders. If a leader always selects the risk solution, it is a risking leader in decisions. And it is conservative if the conservative solution is frequently chosen. For previous bilevel problem, there are some definitions in [1] similar to Definition 0.3 . Actually, when some plan is designed, the followers always report worst cases. Consequently, it is useful for conservative solution. A mean leader will incline to a mean solution. When the Pareto solutions are some discrete points, Definition 0.4 will be changed if necessary. If there are t solutions in Pareto solution set for some x, the definition is

$$x^\star \in \arg\min\{\frac{\sum_{y \in P(x)} F(x, y)}{t}\}.$$

Consequently, three types of solutions are rational and useful in practice.

About the above solutions, there exist some relations to them.

Theorem 0.5 *Denote V_r, V_c and V_m as risk optimal value, conservative optimal value and mean optimal value, respectively. Then, we have*

$$V_r \leq V_m \leq V_c. \tag{7}$$

Proof We show the result by contradiction. Let $\bar{x}_r, \bar{x}_c$ and $\bar{x}_m$ are optimal strategies to risk optimal value, conservative optimal value and mean optimal value, respectively. First, we show $V_r \leq V_c$. We note that

$$V_r = \min_x \min_{y \in P(x)} F(x, y)$$

and

$$V_c = \min_x \max_{y \in P(x)} F(x, y).$$

If $V_r \leq V_c$ were not true, for leader's two types optimal strategies $\bar{x}_r$ and $\bar{x}_c$, there should exist the following formulation

$$min_{\bar{y} \in P(\bar{x}_r)} F(\bar{x}_r, \bar{y}) > max_{\hat{y} \in P(\bar{x}_c)} F(\bar{x}_c, \hat{y}),$$

where $\bar{y}$ and $\hat{y}$ are the corresponding responses of the followers. If we let $\bar{x}_r := \bar{x}_c$, the risk optimal will be no more than V_r. Furthermore, $\bar{x}_c$ is also a feasible strategies of the leader. It is a contradiction. Consequently, $V_r \leq V_c$.

By the similar way, we have the other two parts, namely, $V_m \leq V_c$ and $V_r \leq V_m$. The result is therefore obtained and the proof is complete.

About three types solutions, it seems that it is much more difficult to realize risk optimal solution than others. Among them, if the optimal value is less, it seems more difficult to implement it in practice.

For three types of solutions, we hope to give their optimal conditions, respectively. But it touches the solution of multi-objective solution and it is difficult to describe it. For risk solution, if $(x^\star, y^\star)$ is a risk solution, it is apparent that there should exist the following relations

$$F(x^\star, y^\star) \leq F(x, y)$$

for any $y \in P(x)$ and $G(x, y) \leq 0$.

Similarly, we have the corresponding result to conservative solutions and mean ones.

3. Properties about MPMOC

We first analyze Pareto optimal solution set. Under special cases, Pareto solution set can be described with simple formulations. Certainly, if the Pareto solution set is just a sole point, MPMOC problems are equivalent to previous bilevel programming problems, which can be transformed into a mathematical programming problem with equilibrium constraints. In general, this condition is beyond requirement. Thus, the problem in this paper seems much more difficult than early bilevel problems.

It is fairly difficult to investigate MPMOC in general cases. We therefore consider (1) under special situations. There exists some good properties for convex functions. Accordingly, to obtain some conclusions we give some assumptions as follows:

Assumption 0.6 *(1) Functions g_i, $\quad i = 1, 2, \cdots, m_1$, are all continuously differentiable and convex.*

(2)f_i, $\quad i = 1, 2, \cdots, N$, are all continuously differentiable and strictly convex.

Under Assumption 0.6, lower level multi-objective problems can lead to a single level problem because of the result of Theorem 2 in [2], which is called weighting approach. The weighting method is represented to describe the Pareto solutions to lower level problem by an equivalent problem, which is listed as follows:

$$\begin{array}{ll} \min_y & \sum_{i=1}^N \omega_i f_i(x,y) \\ \text{subject to} & g(x,y) \leq 0, \end{array} \tag{8}$$

where $\omega_i \geq 0$ for $i = 1, 2, \cdots, N$ and $\sum_{i=1}^N \omega_i = 1$.

Obviously, under Assumption 0.6, a point $y^\star$ is the optimal solution of (8) for some x if and only if there exist $\lambda_1^\omega, \lambda_2^\omega, \cdots, \lambda_{m_1}^\omega$ such that

$$\bigtriangledown_y \sum_{i=1}^N \omega_i f_i(x, y^\star) + \sum_{i=1}^{m_1} \bigtriangledown_y g_i(x, y^\star)\lambda_i^\omega = 0, \tag{9}$$

where $0 \leq \lambda_i^\omega \perp g_i \leq 0$ for $i = 1, 2, \cdots, m_1$.

On the other hand, under Assumption 0.6, a point y is a Pareto optimal solution if and only if there exist a sequence $\{\omega_i\}_{i=1}^{m_1}$ such that y is the solution to (8). This is the conclusion of Theorem 2 in [2].

For convenience, we denote

$$\Omega = \{(\omega_1, \omega_2, \cdots, \omega_{m_1})|\omega_i \geq 0, i = 1, 2, \cdots, m_1, \sum_{i=1}^N \omega_i = 1\},$$

and

$$\omega = (\omega_1, \omega_2, \cdots, \omega_{m_1}).$$

Based on the above analysis, we have the following results

Theorem 0.7 *Under Assumption 0.6, $x^\star$ is a risk solution to (1) if and only if $x^\star$ is the optimal solution to the following problem:*

$$\begin{array}{ll} \min_{x,\omega\in\Omega} & F(x, y(x,\omega)) \\ \textit{subject to} & G(x, y(x,\omega)) \leq 0, \end{array} \tag{10}$$

where ω and Ω are given above and $y(x,\omega)$ is the unique solution of the following problem

$$\nabla_y \sum_{i=1}^{N} \omega_i f_i(x,y) + \sum_{i=1}^{m_1} \lambda_i^{\omega} \nabla_y g_i(x,y) = 0, \quad 0 \le \lambda_i^{\omega} \perp g_i(x,y) \le 0, \tag{11}$$

and $y(x,\omega)$ is continuously differentiable.

Proof Firstly, under Assumption 0.6 and for any x, y is a Pareto solution to the lower level problem if and only if there exists a unique vector ω such that y is the solution to (8), which is directly obtained from Theorem 2 of [2].

Under Assumption 0.6, $\sum_{i=1}^{N} \omega_i f_i(x,y)$ is strictly convex for any $\omega \in \Omega$, which is an obvious fact from the definition of strictly convex. Accordingly, if y is the solution to (8) for any x, (11) also holds according to the knowledge of convex analysis (see in [5]).

In summary, for any x, any Pareto optimal solution y to lower level problem corresponds to the solution with some vector ω about (11).

We now show that the existence and uniqueness of $y(x,\omega)$ for any x,ω. By Assumption 0.6, we know that

$$\sum_{i=1}^{N} \omega_i f_i(x,y) + \sum_{i=1}^{m_1} \lambda_i^{\omega} g_i(x,y)$$

is strictly convex. Namely,

$$\nabla_{yy}^2 \sum_{i=1}^{N} \omega_i f_i(x,y) + \sum_{i=1}^{m_1} \lambda_i^{\omega} \nabla_{yy}^2 g_i(x,y)$$

is nonsingular. By implicit function theorem, taking the system of (11) into account, the existence and the uniqueness about solution $y(x,\omega)$ are immediately obtained. Moreover, $y(x,\omega)$ is continuously differentiable.

Finally, from Definition 0.2, (10) is obtained and $y(x,\omega)$ satisfies the conditions in this theorem. This is the conclusion and the proof is complete.

Similarly, for conservative solution and mean solution, we have the corresponding results as follows, respectively. The proof is the same as that of Theorem 0.7 and the detail proof is omitted.

Theorem 0.8 *Under Assumption 0.6, $x^\star$ is a conservative solution to (1) if and only if $x^\star$ is the optimal solution to the following problem:*

$$\begin{array}{ll} \min_x \max_{\omega\in\Omega} & F(x, y(x,\omega)) \\ \textit{subject to} & G(x, y(x,\omega)) \le 0, \end{array} \tag{12}$$

where $y(x,\omega)$ satisfies (11).

Theorem 0.9 *Under Assumption 0.6, $x^\star$ is a mean solution to (1) if and only if $x^\star$ is the optimal solution to the following problem:*

$$\begin{array}{ll} \min_x & \frac{\int_{\omega\in\Omega} F(x,y(x,\omega))d\omega}{\int_{\omega\in\Omega} d\omega} \\ \textit{subject to} & G(x,y(x,\omega)) \leq 0, \end{array} \tag{13}$$

where $y(x,\omega)$ satisfies (11).

(10)-(13) seems very complicated. Problem (12) is min-max one with complementarity constraints. On one hand, complementarity conditions are the constraints of the problems. On the other hand, it is quite hard to cope with min-max problems. It is consequently difficult to cope with them.

But to our fortunate, in the problems the variables ω, x are separable judging from (11). Consequently, they can be attacked with approaches other than those for min-max problems. And there have existed some good results for them, because they are single level optimization problems with complementarity constraints under Assumption 0.6. And many solvers, such as sequential quadratic programming approaches, active set strategies , penalty methods and so on, are all suitable to attack (13).

4. Concluding Remarks

Based on bilevel optimization and practical problems, a new type problem is induced, which is called as mathematical programming with multi-objective constraints (MPMOC) in this paper. MPMOC problems are exceedingly popular in practice. Consequently, it is interesting to investigate them. With some definitions of solutions, we transform MPMOCs into min-max problems and single level problems with equilibrium constraints under certain conditions. Moreover, under general cases, the variables of the problem perhaps are not separable. Therefore, it seems much more difficult to handle them.

MPMOC seems rather difficult. It is a challenging task to attack it. Under certain conditions, MPMOC problems are equivalent to min or min-max problems. Certainly, given strong conditions about F, some further results can be obtained. The detail discussion in this aspect is omitted because such kinds of results have existed. Moreover, the theoretical properties and useful algorithms for MPMOC seem to be another interesting topic.

Actually, when the decisions are uncertain, stochastic approaches are employed. It is very easy to extend mean strategy to stochastic cases. And mean strategy can be regarded as a special stochastic model with equal probability.

Definition 0.10 *We call $u^\star$ a stochastic optimal decision if and only if*

$$x^\star \in \arg\min\{\int_{y\in P(x)} p(y)F(x,y)dy\}, \tag{14}$$

where $p(y)$ is the possibility of y and $\int_{y\in P(x)} dy = 1$.

The stochastic models seem more difficult to solve than deterministic models. The area of stochastic models may be more challenging because it touches some new problems with the appearance of expectation value function. For (14), results about optimal mean decision can be extended to stochastic optimal decision.

As an infant, many properties to MPMOC are unknown till now. It is more crucial to develop some properties to exploit algorithms. On the other aspect, the applications about MPMOC are also interesting.

References

[1] J.F. Bard, Some properties of the bilevel programming problem. *Journal of Optimization Theory and Applications*, **68**(1991), pp. 371-378.

[2] P. Heiskanen P, Decentralized method for computing Pareto solutions in multiparty negotiations. *European Journal of Operational Research*, **117**(1999), pp. 578-590.

[3] Z.Q. Luo, J.S. Pang and D. Ralph, *Mathematical Programs with Equilibrium Constraints*, Cambridge University Press, Cambridge, 1996.

[4] J.S. Pang and M. Fukushima, Complementarity constraint qualifications and simplified B-stationarity conditions for mathematical programs with equilibrium constraints. *Computational Optimization and Applicatins*, **13**(1999), pp. 111-136.

[5] R.T. Rockafellar, *Convex Analysis.* Princeton University Press, Princeton, NJ, 1970.

[6] H. Scheel and S. Scholtes, Mathematical programs with complementarity constraints: stationarity, optimality and sentivity. *Mathematics of Operation Researchs*, **25**(2000), pp. 1-22.

[7] K. Shimizu, Y. Ishizuka and J.F. Bard, *Nondifferentiable and Two Level Mathematical Programming.* Kluwer Academic Publishers, Boston, 1997.

[8] H. von. Stackelberg, *The Theory of the Market Economy.* Oxford University Press, Oxford, 1952.

[9] J.J. Ye, Necessary conditions for bilevel dynamic optimization problems. *SIAM Journal on Control and Optimization*, **33**(1995), pp. 1208-1223.

[10] R. Zhang, Problems of hierarchical optimization in finite dimensions. *SIAM Journal on Optimization*, **4**(1994), pp. 521-536.